2009 IEEE International Conference on Robotics and Automation
Kobe International Conference Center
Kobe, Japan, May 12-17, 2009

A Probabilistic Model for the Performance Analysis of a Distributed Task Allocation Algorithm

Antidio Viguria
Human-Automation Systems Lab
Georgia Institute of Technology
30332 Atlanta, USA
and Center for avanced
Aerospace TECnologies (CATEC)
41300 Seville, Spain
aviguria@catec.aero

Ayanna M. Howard
Human-Automation Systems Lab
School of Electrical
and Computer Engineering
Georgia Institute of Technology
30332 Atlanta, USA
ayanna.howard@ece.gatech.edu

***Abstract*— In this paper we extend our previous work where the mean of the global cost was used as a performance metric for distributed task allocation algorithms. In this case, we move a step forward and calculate the variance of the global cost. This second parameter gives us a better understanding of the distributed algorithm performance, i.e., we can estimate how much the algorithm behavior diverts from its mean. The normal distribution, computed from the theoretical mean and variance, is shown to be suitable for modeling the global cost. This approximation enables us to compare our algorithm theoretically in different cases.**

I. INTRODUCTION

The multi-robot task allocation problem (MRTA) has been studied widely for the last decade. From the different approaches that have been used to solve the general task allocation problem, the distributed approach [3] is considered ideal for teams of robots, and possesses characteristics that fit most robotics applications: high fault tolerance, fast response to dynamic changes in the environment and low computational complexity. We will focus our attention in market-based algorithms [2] since they offer a good compromise between communication requirements and the quality of the solution. They can be considered an intermediate solution between centralized and completely distributed.

We are interested in the Initial Formation Problem [9], a rendition of the general task allocation problem, where each robot can only be allocated to one task. This problem is usually associated with robotic formation control where using local information and control laws, the distributed algorithm is able to drive a given formation error to zero. However, as it is stated in [5], these algorithms require a first step that assigns the robots to the formation positions while taking into account their initial positions, i.e., answer the question who goes where? Also, this problem can be used in exploration missions to allocate the different areas where each robot should take environmental data or look for specific features.

Although the efficiency of market-based algorithms has been evaluated in both simulation [1] and real implementations [4], none of these works has obtained a theoretical bound on the real efficiency of these algorithms. One of the main advantages of theoretical bounds is their capability to compare different algorithms. This is usually a very difficult task since an implementation of the distributed algorithms is needed. An increase of the research work related to theoretical bounds has been occurred in the last years. One of the first results comes from [7] where a bound for a market-based algorithm is calculated. They suppose that all the robots must know all the tasks from the beginning. There are other recent works on theoretical bounds but their algorithms are not based on auctions. In [11] a distributed heuristic with local communication is explained, while in [12] the agents are controlled by hybrid models using distributed potential fields. Both approaches fail to return a highly efficient solution since more than one robot can execute the same task. In [8], a solver of the Traveling Salesman Problem (TSP) is used to decide which robot should execute which task. However, this approach first solves a much more difficult problem to obtain a solution to the assignment problem. All of the commented bound analyses focus on obtaining a worst case bound which is usually very pessimistic and may not ever happen in a real implementation scenario.

In [9], we started a new approach to estimate the behavior of market-based algorithms. Instead of calculating a worst-case metric, where only the worst result is taken into account, we used a probabilistic approach to obtain the expected value (mean) of the global cost, defined as the sum of the costs for all the executed tasks by robots. We consider that this approach is more realistic since considers all the possible results and provides an estimation of the performance over time. This first attempt needed a strong assumption: the cost distribution had to be uniform. We generalized our theory to any kind of cost distribution in [10]. The general results were applied to two different situations and validated with simulations and real experiments.

In this paper, we continue our research work and extend the probabilistic approach calculating the variance of the global cost for one of our distributed algorithms. Our motivation for this work comes from the fact that even though

978-1-4244-2789-5/09/$25.00 ©2009 IEEE

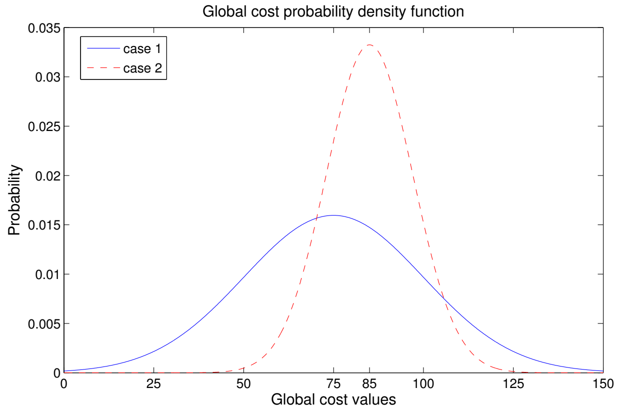

Fig. 1. Probability distributions that model two different global costs. The solid line case has a lower mean but a higher variance than the other case.

the mean is a better metric than that worst-case metric, it is not enough to understand completely the behavior of a task allocation algorithm. For example, in Figure 1 it is shown the distribution of the global cost for two different cases. The solid line case has a lower mean but a higher variance. In this situation, it is not easy to decide which case is the best, i.e., which one has the highest probability of obtaining a lower global cost. In this work, we will show how the mean and the variance can be used to model the global cost by a normal distribution. This will enable us to compute such probabilities, and thus, choose between both situations. As far as we know, this is the first time that the solution (global cost) of a distributed task allocation algorithm is modeled, and an analytical methodology of comparison is explained.

The paper is organized as follows. In next section, we briefly describe the algorithm under study and show how the distributed market-based algorithm is equivalent to a greedy algorithm. In Section III, a probabilistic approach is developed to calculate the variance of the global cost. The obtained theoretical results are applied to the dispersion scenario in Section IV. Then, in Section V, the previous results are extended to situations with different number of robots and tasks. In Section VI, the distribution of the global cost is modeled by a normal distribution and two different scenarios are compared analytically. Finally, conclusions and future work are discussed in Section VII.

II. THE BS-WR ALGORITHM AS A GREEDY ALGORITHM

The BS-WR algorithm is based on a market approach [2] where positions associated with an initial robot formation are recast as biddable tasks in a formation auction. To determine position assignment, a robot agent (auctioneer) dynamically plays the role of announcing the tasks and selecting the lowest cost bid from all the received bids during the auction. Since we focus on robotic formations and tasks will be waypoint tasks, all the costs will be associated with distances. In this algorithm, bidders broadcast their bids only if they do not already have an assigned task, i.e., when a task is allocated to a robot, it no longer bids on other tasks in the auction. This algorithm is easy to implement and uses a small number of messages

We will show that the explained market-based algorithm is equivalent to a centralized greedy algorithm. This does not mean that our algorithm is centralized, only that our distributed algorithm obtains the same solutions as the greedy algorithm. This algorithm is equivalent to the column-scan method [6] for the assignment problem expressed as a matrix where each element is the cost associated with the respective robot and task. We consider that tasks are represented by the columns of the cost matrix and robots by the rows. In this algorithm, each column of the matrix is examined and the row with the lowest cost is selected. The selected row is marked and no longer examined for the rest of the algorithm. Through this process, the algorithm functions as follows:

1) Each column is scanned.
2) The smallest element of the column is selected.
3) The row associated to this element is deleted and not considered for the rest of the algorithm.
4) The same procedure is repeated for the next column until all the columns have been scanned.
5) The selected elements are the solution of the problem and the global cost is the sum of these elements.

An illustrative example will be used to show how both algorithms obtain the same solution.

- The initial positions of the robots and the desired positions of the formation are the ones show in Figure 2.
- The matrix that models this specific problem is:

$$\begin{pmatrix} 30.0 & 41.23 & 20.0 \\ 50.0 & 10.0 & 44.72 \\ 80.0 & 72.11 & 30.0 \end{pmatrix}$$

- Following the algorithm steps, the smallest element of the first column is selected. This element is 30.0 which assigns robot A with task number 1. The row and column of the selected element is deleted and the following matrix is obtained:

$$\begin{pmatrix} 10.0 & 44.72 \\ 72.11 & 30.0 \end{pmatrix}$$

- Next, the smallest element of the second column is selected. This element is 10.0, and therefore, the robot B is assigned to task number 2.
- Finally, the last assignment is made such that robot C is assigned to task number 3. The global cost for this problem is $GC(3) = 70.0$.

As can be observed in Figure 2 and 3, the same tasks are allocated to the same robots using either the BS-WR algorithm or the column-scan method. Therefore, both algorithms are equivalent. The probabilistic approach can be developed using the greedy algorithm, and the results applied to the BS-WR algorithm.

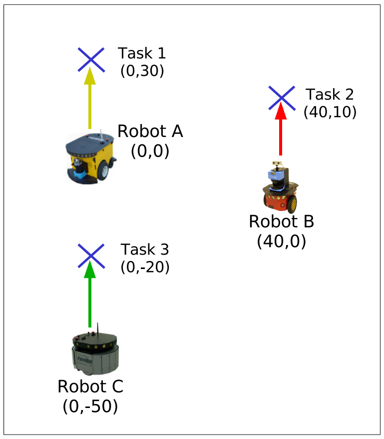

Fig. 2. Initial position of the robots and the desired positions of the formation, and also, the final assignment obtained with the BS-WR algorithm.

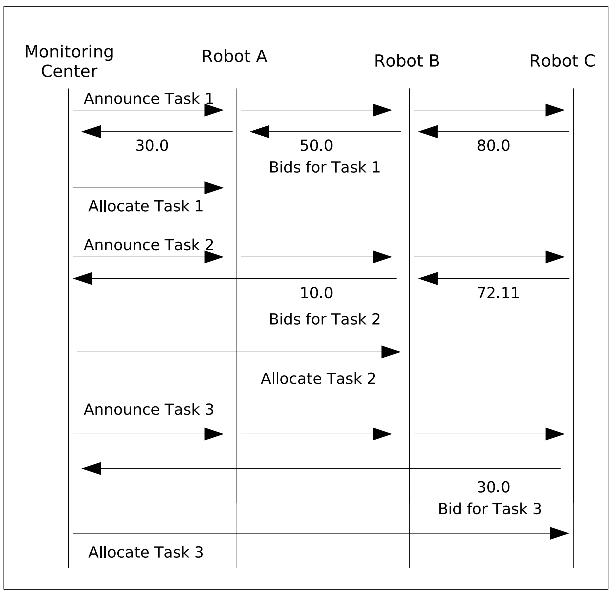

Fig. 3. Messages exchanged in the auction process among the different robots using the BS-WR algorithm. The initial positions of the robots and the positions of the formations are the same as Figure 2.

III. CALCULATION OF THE STANDARD DEVIATION FOR THE BS-WR THE ALGORITHM

The global cost for the BS-WR algorithm is defined as $\sum_{k=1}^{n} m_k$ where m_k is the minimum element of the k^{th} column which has $n-k+1$ elements from the cost matrix, and n is the size of the cost matrix as well as the number of robots and tasks. We define M_k as the minimum of $n-k+1$ independent and equally distributed random variables ($X_{i,k}$) of the k^{th} column, i.e.,

$$M_k \equiv \min\{X_{1,k}, X_{2,k}, X_{3,k}, \ldots, X_{(n-k+1),k}\}.$$

From our previous work [10], we define the expected value of the global cost as

$$E_{GC}(n) = \sum_{k=1}^{n} E(M_k)$$

where $E(M_k)$ is the expected value of each of the costs that come from the executed tasks (minimum elements of each column of the cost matrix) which is calculated as

$$E(M_k) = \int_{-\infty}^{\infty} x \cdot k \left[1 - F_X(x)\right]^{k-1} f_X(x)\, dx \qquad (1)$$

where $F_X(x)$ is the cumulative distribution function and $f_X(x)$ is the probability density function of any of the n^2 random variables that form the cost matrix that models our Initial Formation Problem.

Since the executed tasks are independent (one task allocation does not depend on the already allocated tasks), the variance of the random variable representing the global cost can be calculated as the sum of the variances of the random variables M_k, i.e.,

$$V_{GC}(n) = \sum_{k=1}^{n} V(M_k)$$

where $V(M_k)$ is defined as

$$V(M_k) = E(M_k^2) - E(M_k)^2. \qquad (2)$$

From basic probabilistic theory,

$$E[g(Z)] = \int g(z) f_Z(z)\, dz.$$

Then, $E(M_k^2)$ can be easily calculated as

$$E(M_k^2) = \int_{-\infty}^{\infty} x^2 \cdot k \left[1 - F_X(x)\right]^{k-1} f_X(x)\, dx. \qquad (3)$$

Finally, the variance of the global cost is

$$V_{GC}(n) = \sum_{k=1}^{n} \left[E(M_k^2) - E(M_k)^2\right]$$

where $E(M_k^2)$ is computed from (3) and $E(M_k)^2$ from (1).

IV. APPLICATION TO THE DISPERSION SCENARIO

The dispersion scenario describes a situation where a team of robots are deployed together and afterwards, dispersed around an area. For example, imagine that a team of robots is sent to Mars, and after the landing, they will disperse to explore the area. In this scenario, the costs follow a uniform distribution between $[a, b]$, i.e., $X_{i,j} \sim U(a, b)$.

From [10], we know that the expected value of the global cost for this "uniform" case is

$$E_{GC}(n) = \sum_{k=1}^{n} a + \frac{b-a}{k+1}. \qquad (4)$$

Therefore,

$$E(M_k) = a + \frac{b-a}{k+1}.$$

Also, from (3)

$$E(M_k^2) = \int_a^b x^2 \cdot k \left[1 - \frac{x-a}{b-a}\right]^{k-1} \frac{1}{b-a}\, dx = \frac{k}{(b-a)^k} \int_a^b x^2 (b-x)^{k-1}\, dx.$$

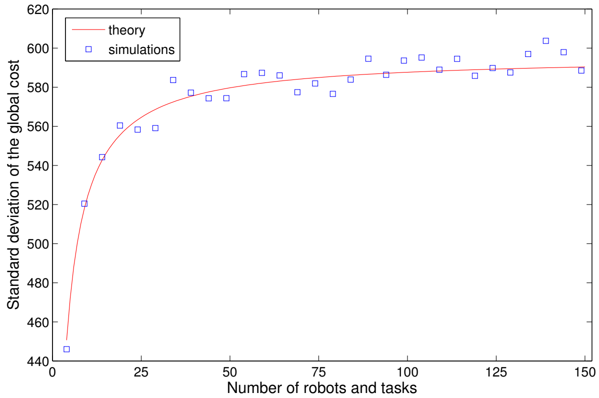

Fig. 4. Standard deviation of the global cost calculated from the theoretical results and simulations. The squares represent the results from simulation applying the BS-WR algorithm over 1000 simulations per number of robots and tasks. The theoretical results, $\sqrt{V_{GC}(n)}$, are shown as a solid line. Both results where calculated in a dispersion scenario with $a = 0$ and $b = 1000$.

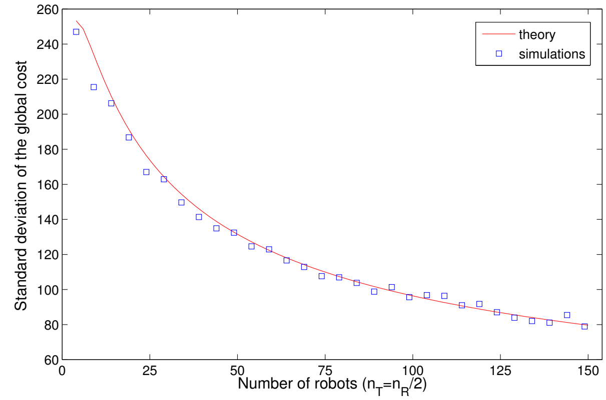

Fig. 5. Standard deviation of the global cost when the number of tasks is half the number of robots ($n_R = n_T/2$). The squares represent the results from simulation applying the BS-WR algorithm over 1000 simulations. The theoretical results, $\sqrt{V_{GC}(n_T, n_r)}$, are shown as a solid line. Both results where calculated in a dispersion scenario with $a = 0$ and $b = 1000$.

Solving the integral by parts where $u = x^2$ and $dv = (b - x)^{k-1}$, it is obtained

$$E(M_k^2) = a^2 + \frac{2\,(a\,(k+1)+b)\,(b-a)}{(k+1)\,(k+2)}.$$

By (2)

$$V(M_k) = \frac{k\,(a-b)^2}{(k+2)(k+1)^2},$$

and thus, the variance of the global cost is

$$V_{GC}(n) = \sum_{k=1}^{n} V(M_k) = \sum_{k=1}^{n} \frac{k\,(a-b)^2}{(k+2)(k+1)^2} \qquad (5)$$

where n is the number of robots and tasks, and a and b are the upper and lower bounds for the uniform distribution that models the costs.

Due to length restrictions, we have only explained how to calculate the variance for the dispersion scenario. However, following similar steps, this approach can be applied to other scenarios, such as the random scenario where the robots and tasks are initially positioned randomly in a square area.

Figure 4 depicts both the theoretical and estimated (from simulations) standard deviation of the global cost. The standard deviation, defined as the square root of the variance, has been used since it makes the visualization of the results easier. The theoretical results have been calculated using (5). For each number of robots and tasks the estimated variance has been computed from 1000 simulations. It can be observed how the simulated and theoretical values are similar.

V. EXTENSION TO DIFFERENT NUMBER OF ROBOTS AND TASKS

In the previous section, the same number of robots and tasks was assumed. Here, we generalize our results for situations when the number of robots (n_R) and tasks (n_T) are different.

A. More robots than tasks

In this case, the cost matrix is not squared and it has more rows than columns. When the allocation is finished, there will be some idle robots. These robots will have the highest costs for the group of tasks.

To compute the variance in this case, it is needed to make a change of variable in (5), $k = n - k' + 1$. Thus,

$$V_{GC}(n) = \sum_{k'=1}^{n} \frac{(n-k'+1)\,(a-b)^2}{(n-k'+3)(n-k'+2)^2} \qquad (6)$$

The column-scan method, which is equivalent to the BS-WR algorithm, scans in this case n_T columns. The k^{th} column has $n_R - k + 1$ elements and there will be $n_R - n_T$ rows or robots at the end of the algorithm without an allocated task. Therefore, (6) changes to

$$V_{GC}(n_T, n_R) = \sum_{k'=1}^{n_T} \frac{(n_R-k'+1)\,(a-b)^2}{(n_R-k'+3)(n_R-k'+2)^2}. \qquad (7)$$

Finally, simulations have been run to evaluate (7). In Figure 5, the theoretical standard deviation is computed using (7) and the simulated standard deviations are calculated from 1000 runs per case (the number of tasks is half the number of robots). It can be observed how the theoretical standard deviation of the global cost almost coincides with the standard deviation obtained from the simulations.

B. Less robots than tasks

The number of columns in the cost matrix is higher than the number of rows. There will be some tasks that will not be executed. These tasks will be the ones that have the highest costs for the group of robots.

The column-scan method only scans the first n_R columns and the k^{th} column will have $n_R - k + 1$ elements. At the end of the algorithm, the last $n_T - n_R$ columns or tasks will not be scanned or allocated to any robot respectively. In this case, it can be shown that (4) changes to

$$V_{GC}(n_R) = \sum_{k=1}^{n_R} \frac{k\,(a-b)^2}{(k+2)(k+1)^2}.$$

It can be observed that V_{GC} does not depend on n_T, since we just allocate the tasks in order. When there are no more robots left, we just discard the rest of the tasks. Therefore, the problem is equivalent to having n_R number of robots and tasks.

VI. NORMAL APPROXIMATION OF THE GLOBAL COST DISTRIBUTION

In the previous sections, we have shown how to compute the mean (expected value) and variance of the global cost distribution, when the Initial Formation Problem is solved using the BS-WR algorithm. Even though these parameters provide certain information about the distribution, they do not completely characterize it. Since the real CDF (cumulative distribution function) is hard to calculate analytically, we used the information provided by the mean and variance to obtain an approximation by a normal distribution. The normal distribution is well-known and has shown to be suitable to model very different situations. In addition, the Kolmogorov-Smirnov test was used to validate this choice.

The procedure to be followed is the next one. First, we calculate the mean and variance of the global cost for a specific case (such as the distribution of the matrix costs, and the number of robots and tasks) using our probabilistic approach. Since the normal distribution is completely described by its mean and variance, we use the previously computed parameters as those of the normal distribution.

In order to test this methodology, we have simulated the BS-WR algorithm for 10 robots and tasks with costs uniformly distributed between 0 and 1000. In Figure 6, it can be observed how the CDF of the real distribution (empirically calculated from 10000 simulations) is very similar to the CDF of a normal distribution calculated from the mean (4) and variance (5) formulae that have been obtained in our probabilistic approach.

Once we have approximated the global cost distribution by a normal distribution, we can compare analytically different cases and study which is the probability that one case obtains a lower global cost. For example, supposing that we have applied our BS-WR algorithm to two different scenarios, and obtained the approximated distributions that are shown in Figure 1. It is very difficult to decide which is the best situation, since one distribution has a lower mean but also a higher variance. However, this can be decided theoretically calculating which distribution has the highest probability to obtain the lower value (global cost). In general, the two random variables that approximate the global cost distribution in these two scenarios will be denoted by X, for scenario number 1, and Y for the scenario number 2. The probability that the random variable X is lower than Y is

$$P(X<Y) = \int_{-\infty}^{\infty} P(X \leq Y|y) \cdot f_Y(y)dy.$$

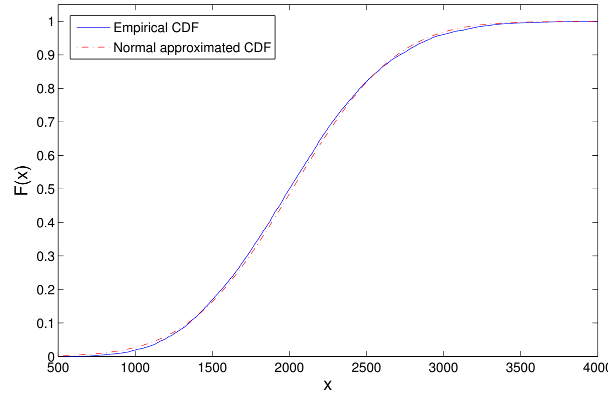

Fig. 6. Comparison of the empirical CDF and the normal approximation whose parameters are computed using our probabilistic approach. The empirical CDF has been calculated from 10000 simulations using the BS-WR algorithm in a 1000mx1000m arena with the costs uniformly distributed between 0 and 1000.

or equivalently

$$P(X<Y) = \int_{-\infty}^{\infty} F_X(y) \cdot f_Y(y)dy. \quad (8)$$

Thus, $P(X<Y)$ represents the probability that the first global cost X will obtain a lower global cost than in the second case. Also, this probability provides information about how much better one algorithm is in comparison with the other one.

The previous formula can be extended to compare N number of cases. In this situation, the global costs are approximated by the probabilistic distributions $X_1, X_2, X_3, \ldots, X_N$. The probability that the global cost X_i will be lower than the rest of the global costs is

$$P(X_i < X_{-i})$$

where $X_{-i} = \{X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_N\}$. Assuming that all the global costs are independent, then

$$P(X_i < X_{-i}) = \prod_{j \neq i} P(X_i < X_j),$$

where each of these probabilities has the same form as in (8).

This probability, $P(X_i < X_{-i})$, is calculated for all the cases under study, and the one with the highest probability is chosen, i.e., the best case will be X_i iff

$$P(X_i < X_{-i}) > P(X_j < X_{-j}) \quad \forall j \neq i.$$

We have applied the BS-WR algorithm to two different scenarios with 10 robots: a dispersion scenario with the costs following a uniform distribution $[0, 1000]$, and a random scenario (where the robots and tasks are initially positioned randomly in a square area) with a square area of 707.1 units per side. Therefore, in both cases, the costs range from 0 to 1000 units but with different distributions. Then, the mean and the variance are calculated using our probabilistic approach:

- Dispersion scenario (X): $\mu_1 = 2019.9$ and $\sigma_1^2 = 524.69^2$.
- Random scenario (Y): $\mu_2 = 2174.1$ and $\sigma_2^2 = 484.13^2$.

These values are used to approximate the global costs distributions using a normal distribution. Afterwards, we compare both cases analytically and calculate $P(X < Y)$ which is equal to 0.5854. Therefore, the BS-WR algorithm has a highest probability to obtain a lower global cost in the dispersion scenario. Moreover, the BS-WR algorithm, applied to the dispersion scenario, will obtain in 58.54% of the experiments better results than in the random scenario. It is important to point out that this percentage has been calculated without running a single simulation.

Finally, the BS-WR algorithm has been simulated for 10000 experiments for both scenarios. It has been verified that the BS-WR algorithm obtains better results for the dispersion scenario in 58.16% of the experiments which is very similar to the theoretical value. Therefore, this approach can be used to compare analytically the same algorithm in different scenarios or different algorithms without the need of performing thousands of simulations.

VII. CONCLUSIONS AND FUTURE WORK

A complete probabilistic study for the BS-WR algorithm has been developed. First, we have calculated the variance of the global cost distribution. Then, we have applied this result to the dispersion scenario. This enables us to obtain a better knowledge of the algorithm behavior. The variance provides information about how much the algorithm diverts from its mean. The theoretical formula of the variance has been validated with simulated data.

Since the real distribution of the global cost is hard to obtain analytically, a normal distribution has been chosen for its approximation. The theoretical mean and variance have been used to describe the normal distribution. Once we have modeled the random variable of the global cost with a normal distribution, an analytical procedure has been explained to compare two or more cases. These cases can be one algorithm in different scenarios, or different algorithms in the same scenario. The theoretical procedure has been validated with the BS-WR algorithm applied to two scenarios (dispersion and random) with very similar results between the simulations and the theoretical values. As far as we know, these are the first steps that deal with the performance comparison of a distributed algorithm in different scenarios without the use of simulations.

In our future work, we plan to extend this probabilistic framework to other distributed task allocation algorithms that solve the Initial Formation Problem. This framework will enable us to compare different algorithms in any scenario without the need of simulations.

ACKNOWLEDGMENTS

This work supports research in Reconfigurable Sensor Networks for the National Aeronautics and Space Administration, Earth Science Technology Office. The first author would like to thank the Fulbright commission in Spain for providing funding through a Fulbright scholarship.

REFERENCES

[1] M. B. Dias and A. Stenz. Opportunistic optimization for market-based multirobot control. In *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2714–2720, Lausanne, Switzerland, 2002.

[2] M.B. Dias, R. Zlot, N. Karla, and A. Stentz. Market-based multirobot coordination: A survey and analysis. *Proceedings of the IEEE*, 94(7):1257–1270, 2006.

[3] B.P. Gerkey and M. J. Matarić. Murdoch: Publish/subscribe task allocation for heterogeneous agents. In *Proceedings of the Fourth International Conference on Autonomous Agents*, pages 203–204, Barcelona, Spain, 2000.

[4] B.P. Gerkey and M. J. Matarić. Sold!: Auction methods for multi-robot coordination. *IEEE Transactions on Robotics and Automation, Special Issue on Multi-robot Systems*, 18(5):758–768, 2002.

[5] M. Ji and M. Egerstedt. Role-assignment in multi-agent coordination. *International Journal of Assistive Robotics and Mechatronics*, 7(1):32–40, 2006.

[6] J. M. Kurtzberg. On approximation methods for the assigment problem. *Journal of the ACM*, 9(4):419–439, 1962.

[7] M.G. Lagoudakis, M. Berhault, S. Koenig, P. Keskinocak, and A.J. Kleywegt. Simple auctions with performance guarantees for multi-robot task allocation. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, volume 1, pages 698–705, Sendai, Japan, 2004.

[8] S. L. Smith and F. Bullo. Target assignment for robotic networks: asymptotic performance under limited communication. In *Proceedings of the American Control Conference*, pages 3585–3590, New York, USA, 2007.

[9] A. Viguria and A. Howard. Upper-bound cost analysis of a market-based algorithm applied to the initial formation problem. In *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 2326–2331, San Diego, USA, 2007.

[10] A. Viguria and A. M. Howard. Probabilistic analysis of time-distributed algorithms for initial robotic formations. *IEEE Transactions on Robotics*. Submitted May 2008.

[11] S. Yun and D. Rus. Optimal distributed planning of multi-robot placement on a 3d truss. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1365–1370, San Diego, USA, 2007.

[12] M. M. Zavlanos and G. J. Pappas. Dynamic assignment in distributed motion planning with limited information. In *Proceedings of the American Control Conference*, pages 1173–1178, New York, USA, 2007.